



## King's Research Portal

DOI:

[10.22618/TP.PCMS.20164.349030](https://doi.org/10.22618/TP.PCMS.20164.349030)

*Document Version*

Publisher's PDF, also known as Version of record

[Link to publication record in King's Research Portal](#)

*Citation for published version (APA):*

Sarkadi, S. (2016). Artificial Consciousness in an Artificial World. In *Communication Today: An Overview from Online Journalism to Applied Philosophy* (pp. 322-330). Trivent Publishing.  
<https://doi.org/10.22618/TP.PCMS.20164.349030>

### **Citing this paper**

Please note that where the full-text provided on King's Research Portal is the Author Accepted Manuscript or Post-Print version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version for pagination, volume/issue, and date of publication details. And where the final published version is provided on the Research Portal, if citing you are again advised to check the publisher's website for any subsequent corrections.

### **General rights**

Copyright and moral rights for the publications made accessible in the Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the Research Portal

### **Take down policy**

If you believe that this document breaches copyright please contact [librarypure@kcl.ac.uk](mailto:librarypure@kcl.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.

Trivent Publishing

© The Authors, 2016

Available online at <http://trivent-publishing.eu/>



Series: *Philosophy, Communication, Media Sciences*

Volume: *Communication Today: An Overview from Online Journalism to Applied Philosophy*

# Artificial Consciousness in an Artificial World

Ştefan Sarkadi

King's College London, Department of Informatics, United Kingdom, [stefansarkadi@gmail.com](mailto:stefansarkadi@gmail.com)

## Abstract

*This paper presents the idea that using computers to simulate complex virtual environments can give rise to artificial consciousness within those environments. Currently, limitations in creating artificial consciousness may be imposed by material compounds that enable the transmission of signals through artificial systems such as robots. Virtual environments, on the other hand, provide the necessary tools for surpassing such limitations. I also argue that using virtual reality tools to enable complex interaction between humans and AI (artificial intelligence) within virtual environments is the most reasonable way to obtain artificial consciousness by appealing to the nature of human bias.*

## Keywords

Artificial Intelligence; Artificial Consciousness; Virtual Reality; Virtual Consciousness; Human-Computer Interaction; Computational Modeling of Consciousness; Artificial Life; BayesAct.

This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC-BY-NC-ND 4.0) license, which permits others to copy or share the article, provided original work is properly cited and that this is not done for commercial purposes. Users may not remix, transform, or build upon the material and may not distribute the modified material (<http://creativecommons.org/licenses/by-nc/4.0/>)

DOI: 10.22618/TP.PCMS.20164.349030

## I. Introduction

Before reading further on, the reader must understand what this article is and what it is not. Let us start by saying what it is not, so that I can eliminate any bias towards the ideas that will follow later in the paper. I shall even dare to postulate from the beginning that consciousness is information processing given the fact that I adopt the computationalist perspective on what is known as cognition.<sup>1</sup>

This is not a philosophical paper. Even though it tackles philosophical problems, especially issues regarding philosophy of mind and epistemology, the paper's aim is not to solve these problems, but its aim is to offer a new approach that will shed light on some of these problems.

This is not a pure computer science paper, because it does not offer any rigorous mathematical methodology or formulas to be applied in the design of virtual worlds or artificial consciousness. However, it presents conceptual frameworks and methods that can be used to take current research regarding artificial consciousness and virtual reality (henceforth VR) a step further.

This is not an orthodox cognitive science paper, because it does not use certain tools used in cognitive science to study the human mind or to develop or analyze theories about the human mind. It can be considered a cognitive science paper in respect of the relevant knowledge it derives from cognitive science to build a new perspective on artificial consciousness<sup>2</sup> and artificial life.

The article is about a unifying perspective on AI<sup>3</sup> and artificial consciousness. It presents various ideas and methods in cognitive science, linguistics and computer science in order to speculate how consciousness might arise from virtual environments such as computer games. Ideally, this paper can be described as a type of research proposal in human computer interaction<sup>4</sup> that focuses on the emergence of artificial consciousness.

Globally, artificial consciousness is a very sensitive issue and it becomes even more sensitive within philosophy of mind and computer science. This is mainly due to the problematic nature of consciousness in general, and the limitations of computability in particular. In this paper I will not address issues such as The Hard Problem or anything related to the phenomenal aspect of consciousness.<sup>5</sup> The only exception from this rule are indirect references to the basic notion of subjective experiences.<sup>6</sup> Instead I will focus on the human bias towards on the possibility of artificial consciousness and the possibilities of attaining artificial consciousness through VR.<sup>7</sup>

## II. Virtual Environments

First, I want to explain some assumptions I make before presenting the idea that virtual environments serve as a necessary condition for the rise of artificial consciousness and how together with several other conditions may be sufficient for artificial consciousness to emerge.<sup>8</sup>

The first assumption I make is that the performance of today's hardware technologies<sup>9</sup> are insufficiently complex to serve a sufficiently complex transmission of information.<sup>10</sup> This is a crucial

---

<sup>1</sup> Dietrich, Eric. "Computationalism." *Social Epistemology: A Journal of Knowledge, Culture and Policy* 4, no. 2 (1990): 135-154.

<sup>2</sup> Chrisley, Ron. "Philosophical foundations of artificial consciousness." *Artificial intelligence in Medicine* 44, no. 2 (2008): 119-137.

<sup>3</sup> Negnevitsky, Michael. *Artificial intelligence: a guide to intelligent systems*. Pearson Education, 2005.

<sup>4</sup> Helander, Martin G., ed. *Handbook of human-computer interaction*. Elsevier, 2014.

<sup>5</sup> Chalmers, David. "The hard problem of consciousness." *The Blackwell companion to consciousness* (2007): 225-235.

<sup>6</sup> Jackson, Frank. "Epiphenomenal qualia." *The Philosophical Quarterly* 32, no. 127 (1982): 127-136.

<sup>7</sup> Sloman, Aaron, and Ron Chrisley. "Virtual machines and consciousness." *Journal of Consciousness Studies* 10, no. 4-5 (2003): 133-172.

<sup>8</sup> Blommestijn, Suzan Quirine. "Emergence and Consciousness.", <http://dare.uva.nl/cgi/arno/show.cgi?fid=511797> (accessed December 2015).

issue if we also accept that information processing underlies the nature and complexity of a conscious agent.<sup>11</sup>

How can VR eliminate such material and technological limitations? In order to answer this question we must look at what VR consists of. Even though any VR platform runs on a certain type of hardware—which is itself made out of certain materials with certain physical properties—VR is not limited to the physical hardware that it supervenes upon. VR is further dependent on the software that runs on (i.e., is enabled by) the underlying hardware, because on the basis of this programmable software, we can build virtual worlds whose virtual properties do not have to be identical to the physical properties of the underlying hardware. In this way it is possible, in principle, to simulate the properties of any possible world, with the only technological limitation being that the underlying hardware is sufficiently programmable.

Through programmable software, we were able to build worlds. Let us think of computer games for example. If you have been an avid gamer since the 1980s or '90s, you can recall how fast the world of gaming has developed. Let us take the Grand Theft Auto series, for instance. I specifically chose Grand Theft Auto (henceforth GTA) because of its gameplay. The gameplay in GTA is described as open-world gameplay. This means there is a very low degree of limitation in respect to what your character can do when playing the game. For example, you can use your character to roam freely through the virtual world of the game. You can drive cars, motorcycles, boats, choose to start various missions assigned to you by different characters in the game etc. You can do all of this while following a well defined storyline that takes place in a virtual world.<sup>12</sup>

In contrast to a robot that has a body made out of various physical materials like metal, plastic and silicon boards, a virtual agent's body is defined by the computations on a silicon chipboard. Therefore, the properties of the virtual agent's body can be manipulated in different ways in order to achieve greater performance. For example, increasing or decreasing the density of virtual nerve cells on a virtual skin in order for the virtual agent to receive more or less information from the virtual environment.

Let us take iCub, for instance.<sup>13</sup> iCub is a humanoid robot that has been designed using theories from embodied cognition. iCub has been developed to aid research in cognitive science because it permits scientists to test various hypotheses about cognition. Hypotheses about language acquisition, for instance, have been successfully tested by research teams that used iCub in their experiments. However, even though iCub might arbitrarily be the most advanced humanoid robot up to date it fails to interact convincingly with the environment it is put in for us to consider it as intelligent as a human child. Not to mention having a consciousness, which is even more problematic.

The software engine behind iCub captures information from the physical environment and digitizes it in order for the tiny robot to process the information. The software that translates information is highly dependent on the physical architecture of the robot. The physical architecture of iCub is very robust and simplistic compared to the organic architecture of a human. The ongoing physical processes that iCub's body suffers are extremely scarce if we are to compare it to all the ongoing processes in the human body. Where are all the chemical reactions? Why does the iCub's processor

---

<sup>9</sup> Millán, José del R., Rüdiger Rupp, Gernot Mueller-Putz, Roderick Murray-Smith, Claudio Giugliemma, Michael Tangermann, Carmen Vidaurre et al. "Combining brain-computer interfaces and assistive technologies: state-of-the-art and challenges." *Frontiers in neuroscience* 4 (2010): 161.

<sup>10</sup> Terzopoulos, Demetri. "Artificial life for computer graphics." *Communications of the ACM* 42, no. 8 (1999): 32-42.

<sup>11</sup> Chalmers, David J. "How can we construct a science of consciousness." *The cognitive neurosciences III* (2004): 1111-1119.

<sup>12</sup> North, Rockstar, Wydawca Rockstar Games, and Wydawca Cenega Poland. "Grand theft auto." *Rockstar Games* (1998).

<sup>13</sup> Tsagarakis, Nikolaos G., Giorgio Metta, Giulio Sandini, David Vernon, Ricardo Beira, Francesco Becchi, Ludovic Righetti et al. "iCub: the design and realization of an open humanoid platform for cognitive and neuroscience research." *Advanced Robotics* 21, no. 10 (2007): 1151-1175.

not release equivalent information to the way the brain does when releasing dopamine? Why does the colour of iCub's artificial cheeks not change in correlation to some physical change in the transmission of electric signals equivalent to blood flow in a human counterpart. So on and so forth.

These complex processes that I mentioned in the paragraph above are not only to be implemented only for a human to believe the robot is genuinely experiencing something, but they should be present as functions. When I say functions, I refer to the functional role of these processes within an artificial or virtual organism. That is, their role should not be reduced to one of signal transmission from social agent A to social agent B. Why is that so? Because if it is only reduced to transmission of signals to be interpreted in a social environment, then the phenomenon itself can be further reduced to language simulation. Language simulation is not enough. We know this from current ongoing research. However, some principles behind it are relevant. I will explain relevance of these principles in the following part of the paper.

### III. Computational Modeling

What is computational modeling?<sup>14</sup> Well, as the name suggests, it is a type of modeling used to represent something. Firstly, it models a phenomenon or a system. In the second place, it models something using computations. Usually, a computational model is a representation of a physical phenomenon/systems. Hence, I need to make the distinction between the physical and the computational instantiation of a given phenomenon or system.

In science, modeling is used to observe and study phenomena/systems in a controlled environment. In the early stages of empirical science (more often in experimental physics than other empirical sciences), modeling was done using physical models. A physical model is a reproduction of observable phenomena/systems at a scale that can be controlled and manipulated. Having control over the phenomena/systems using a model means that the researcher or scientists can manipulate various parameters.

Using computational models to study complex phenomena or complex systems makes modeling simpler and much more efficient.<sup>15</sup> Running simulations on a computer means that less resources are needed for building the model, and less time is needed for playing with the parameters and getting the results. In other words computational modeling is cheaper, faster and better than classical modeling. But what does computational modeling have to do with Artificial Consciousness? Or what is the reason I decided to talk about computational modeling in the first place?

Well, if artificial consciousness can arise from virtual or simulated environments, one must provide the necessary tools to study artificial consciousness within these environments. Given the fact that these environments are virtual and simulated, it is reasonable to assume that computational modeling is the best way to study phenomena/systems within these environments. That is because computational modeling is a simulated virtual environment in itself. Our future aim is to create extremely complex systems that would enable Artificial Consciousness to emerge. Classical modeling in this case would require resources that current technological development cannot provide yet. Not to mention the costs of materials and work that would need to be done to model such environments on a physical level and control them.

Let us take a particular example of using computer simulation to study a given system or phenomenon. In linguistics, or evolutionary linguistics to be more precise, computer simulations are being used to study the evolution of language in a simulated space and time.<sup>16</sup> In order to simulate language, researchers in computational linguistics define virtual agents that communicate with each

---

<sup>14</sup> Bryden, John, and Jason Noble. "Computational modeling, explicit mathematical treatments, and scientific explanation." (2006): 520-526, <http://eprints.soton.ac.uk/265242/1/brydenCompModelling.pdf> (accessed December 2015)

<sup>15</sup> Edmonds, Bruce. "Complexity and scientific modeling." *Foundations of Science* 5, no. 3 (2000): 379-390.

<sup>16</sup> Kirby, Simon, Tom Griffiths, and Kenny Smith. "Iterated learning and the evolution of language." *Current opinion in neurobiology* 28 (2014): 108-114.

other using a predefined language structure. Parameters are also defined in order to create some sort of environment in which the agents interact in a certain way using language. This method is probably derived from the idea of cellular automata. Conway's Game of Life is a good exemplification of this process.<sup>17</sup>

A cellular automaton is basically a two dimensional grid of cells that change states. Each cell interacts with its eight neighbours. This interaction is defined by a set of rules. This set of rules can be a minimalistic one, or a very complex one. An example of a minimalist set of rules is the one in the Game of Life. The Game of Life has four rules:

1. Any living cell with fewer than two live neighbours dies - the under-population rule.
2. Any living cell with two or three live neighbours lives on to the next generation.
3. Any living cell with more than three live neighbours dies - the over-population rule.
4. Any dead cell with exactly three live neighbours becomes a live cell - the reproduction rule.

Being simultaneously applied to every cell on the two dimensional grid, these four rules determine an initial pattern that is called the "seed" of the system, which gives rise to the first generation of cells. After several generations, more patterns emerge, suffering changes in complexity. This phenomenon occurs without bringing any change to the initial rules.

The major conceptual difference between the cellular automata in the Game of Life and current language simulators is the concept of language itself. The agents in these simulations communicate meaning. Not complex meanings, but very basic ones—so basic that they are reduced to different binary combinations of 1s and 0s within matrices. As I mentioned in the previous part of the paper, language simulation is not enough.

In a more recent study<sup>18</sup> done by Hoey, Schröder and Alhothali, they have managed to simulate a way in which artificial agents interact taking into account emotions. However, the emotions in this study are not themselves simulated. The basis of their findings relies on how agents behave linguistically when they take into account certain biases about emotions. This model is called the Bayes Act. On one hand, the Bayes Act has proven to be a very useful "plug-in" tool for artificial intelligent systems when interacting with humans. This can be considered a huge leap in human-computer interaction. On the other hand, the model itself is still just a shortcut towards attaining a limited linguistic behaviour. An argument against this type of approach is that this is a simulation of behaviour given biases of subjective experiences that lacks the simulation of subjective experiences themselves. If we want a system to be considered conscious, then it would be reasonable to include the simulation of conscious experiences as well.

Another reason why language simulation is not enough is due to the fact that in order to simulate language using computational modeling, one does not need to define complex meanings and interactions and environments. This entails that there is a huge gap in the degree of complexity between environments in which consciousness arises. Such as the difference between the current world we live in and environments created by computational linguists. In other words, these artificial environments lack complexity. The same reason can be applied not only to the environments themselves, but to the agents within those environments. They simply are not complex enough.

Why is it not enough? From a psychological perspective on bias, they are not complex enough for a human individual to believe that consciousness is present. In order for a system/agent to be considered conscious by a human, it must meet certain requirements. If these requirements are met, then it is reasonable for a human to assume that the system is conscious. We will see what these requirements are in the next part of the paper.

<sup>17</sup> Conway, John. "The game of life." *Scientific American* 223, no. 4 (1970): 4.

<sup>18</sup> Hoey, Jesse, Tobias Schröder, and Areej Alhothali. "Affect control processes: Intelligent affective interaction using a partially observable Markov decision process." *arXiv preprint arXiv:1306.5279* (2013).

#### IV. Determining Artificial Consciousness

I believe that the most important aspect in determining consciousness is determining the degree of social interaction. If one manages to achieve a virtual simulation of social interaction, then it means one has managed to simulate consciousness. Assigning consciousness is therefore dependent on some sort of activity of agents in social interactions.<sup>19</sup>

The dependency of consciousness on social interaction is primarily due to the fact that consciousness is assigned by social agents to other social agents. Similar to properties relating to ethical behaviour such as being a good or a bad person, one can be a person with or without consciousness. In the case of consciousness, it is the complexity of behaviour that determines one agent to assign consciousness to another, and not the type of behaviour as in the case of ethical properties.

One might even postulate a formalization of requirements to be met in order to be considered conscious. However, this is a trickier task than one might initially think. It is difficult to assess all the behavioural and physiological cues in order to check the requirements of an agent to be considered conscious. Such cues include, but are definitely not limited to: sensory motor coordination of the agent when interacting with the environment (body movement when walking, talking etc.), linguistic behaviour (words used, syntax, phrasing, pronunciation etc.), somatic responses (sweating, shuddering, crying, vomiting etc.). These types of cues are usually assessed by agents when assigning consciousness to other agents whether they are aware of the process or not, going beyond basic linguistic behaviour.

The fact that some agents are unaware of this process may be argued through the cognitive evolution framework. The human brain has developed in such a way as to not use too much processing power for this particular task. Thousands of years of genetic development have resulted in this algorithmic efficiency.<sup>20</sup> Therefore, today we have become unaware of this sub-process.

It is reasonable then to say that there is a strong prior bias against assigning consciousness to artificial social agents such as robots. In other words, we do not find it “natural” for robots to have consciousness because they have not emerged similarly to organisms like us, humans.<sup>21</sup>

This is why I tend to believe that one way of solving the problem of artificial consciousness is by addressing to this human bias instead of focusing on the development of materials, hardware, or even software in order to create the ultimate conscious robot. The argument that I bring is the following: even if we reach the limit of technological development, this bias may still remain. Therefore, a robot can exhaustively exhibit all possible behavioural and physiological human traits, but will not be assigned consciousness due to the fact that it is simply just an artificial social agent, and not a “real,” organic, human, social agent.

There are two main procedures I suggest to be applied in order to eliminate this bias towards artificial social agents. The first procedure is to develop complex artificial agents, which manage to exhibit a behaviour that is highly similar to humans’ behaviour. Second, is to make humans interact with the artificial agents in a way in which they will be determined by the complexity of the robots’ behaviour to not discriminate between robots and other humans.

These procedures do not offer an absolute solution to the artificial consciousness problem. However, they offer a sufficient condition for humans to assign artificial consciousness to robots. Of course, this does not mean that all humans will start assigning consciousness. Some humans might just

<sup>19</sup> Schweizer, Paul. "Could There be a Turing Test for Qualia?." *Revisiting Turing and His Test: Comprehensiveness, Qualia, and The Real World*: 41, <http://www.frontiersinai.com/turingfiles/July/07.pdf#page=41> (accessed December 2015)

<sup>20</sup> Geary, David C., and Kelly J. Huffman. "Brain and cognitive evolution: forms of modularity and functions of mind." *Psychological bulletin* 128, no. 5 (2002): 667.

<sup>21</sup> Evans, Jonathan St BT. *Bias in human reasoning: Causes and consequences*. Lawrence Erlbaum Associates, Inc, 1989.

have too strong biases against artificial consciousness and will never agree that what they believe to be an artificial agent will ever be conscious.

An interesting phenomenon might arise from these strong biases. Humans with these types of biases might confront themselves with uncertainty. Which can be explained through the following line of reasoning: if these humans are to be tricked into thinking the artificial social agent is human, they can just as well be tricked that the human social agent is a robot. That is because all the data that they can empirically verify is the behaviour. Therefore, these humans will never know if they are interacting with a robot or with a human. This might result in some sort of paranoid/discriminative human agents that will constantly try to test for consciousness in every social agent they interact with.

## V. Attaining Complexity and Changing Bias through VR

For the majority of human agents, super strong biases are not an issue that blocks the assignment of artificial consciousness. However, the bias against assigning artificial consciousness still remains. So, what is the best method to attain complexity and to eliminate this bias? The answer we are looking for might come from a technology that is already available to us. The reason I talked about virtual environments in the beginning of the paper was to create a framework for introducing VR tools/hardware in this last part.

When I discuss VR tools I mainly refer to instruments such as the Oculus Rift, Data Gloves, 3D Stereoscopes etc. These tools enable a given human individual to immerse into a virtual environment (mostly computer generated game worlds) by establishing sensory data bridges between the virtual environment and the human body. This technology has already been extremely successful in treating patients with sensory motor difficulties, for example. This entails the fact that the human mind/brain and physiology is highly responsive to virtual information. Or maybe just as responsive as it is to non-virtual information. For example skin tone bias in VR is consistent with skin tone bias in real life.<sup>22</sup>

If the physiology and behaviour of the human body is responsive to information conveyed by a computer simulation, then it is reasonable to assume that in principle we can get a multitude of responses using this method. Responses such as the ones when humans interact with other humans in a social environment. Some of these responses are triggered by certain types of information like cues from other social agents. We briefly talked about the limitations of iCub to convey functional physiological cues (blushing, crying etc.) that are very similar to a human's, which is due to material limitations.

A virtual iCub would not face these limitations. In principle, a virtual iCub can be built in such a way as to convey all the necessary informational cues in order for it to interact with humans in a sufficiently complex way as to trigger emotions. Ideally, this relationship between human and artificial agent would be a two way channel for information transmission. This means that the virtual iCub would also be influenced by a type of virtual avatar of the human, which mirrors all of the human's physiological responses and behaviour. Therefore, human and artificial agent will be able to exchange information through a solid channel.

The ideal way for such an interaction to take place would be a computer game based platform. One of the reasons behind this is that humans tend to emotionally engage in this sort of activity, which might considerably increase the bias towards assigning the virtual iCub conscious experience, thus eliminating the bias against assigning consciousness to artificial agents. A game world would offer both agents various types of activities to engage in freely. The greater and more complex the game world, the greater and more complex the interaction between artificial and human agent, which means that a change in bias is more rapid.

---

<sup>22</sup> Rossen, Brent, Kyle Johnsen, Adeline Deladisma, Scott Lind, and Benjamin Lok. "Virtual humans elicit skin-tone bias consistent with real-world skin-tone biases." In *International Workshop on Intelligent Virtual Agents*, (Berlin: Springer Berlin, 2008), 237-244.



Finally, we can say that there have to be two main variables for complexity. The first one is the potential physiological and behavioural complexity of virtual agents. And the second one is the complexity of the virtual environment in which the interaction takes place.

## VI. Conclusion

Given the ideas presented in this paper together with the fact that emotional responses and biases in humans have been consistently studied in cognitive science, my proposal is to research how to methodically achieve a change in bias towards accepting consciousness in machines by getting repeated instantiations of certain types of somatic and sociologic responses in humans. In this way, humans will anthropomorphize virtual agents such as a virtual iCub. Hence, they will assign consciousness to the virtual agent. In other words, get strong correlations of emotional responses from humans when they interact as virtual agents with their artificial counterparts. This entails the fact that the responses are determined by a certain belief system which resulted from repeated exposure to a certain type of data extracted from virtual artificial agents. In conclusion, due to the social nature of consciousness, it is more reasonable to obtain artificial consciousness in the near future by working on human psychology rather than by just focusing on the construction of complex hardware systems.

The applications that follow from this type of study are numerous. They may range from game design and virtual or semi-virtual web platforms to cognitive-behavioural treatment of patients. These applications point towards a change in the paradigm in which modern societies work. Here, ethical implications of cyber-enhancements, A.I. and VR are huge. However, care must be taken in the implementation of A.I. as pointed out by numerous researchers and specialists.<sup>23</sup>

## Bibliography

- Blommestijn, Suzan Quirine. "Emergence and Consciousness." <http://dare.uva.nl/cgi/arno/show.cgi?fid=511797> (accessed December 2015).
- Bryden, John, and Jason Noble. "Computational modeling, explicit mathematical treatments, and scientific explanation." (2006): 520-526, <http://eprints.soton.ac.uk/265242/1/brydenCompModelling.pdf> (accessed December 2015)
- Chalmers, David J. "How can we construct a science of consciousness." *The cognitive neurosciences III* (2004): 1111-1119.
- Chalmers, David. "The hard problem of consciousness." *The Blackwell companion to consciousness* (2007): 225-235.
- Chrisley, Ron. "Philosophical foundations of artificial consciousness." *Artificial intelligence in Medicine* 44, no. 2 (2008): 119-137.
- Conway, John. "The game of life." *Scientific American* 223, no. 4 (1970): 4.
- Dietrich, Eric. "Computationalism." *Social Epistemology: A Journal of Knowledge, Culture and Policy* 4, no. 2 (1990): 135-154.
- Edmonds, Bruce. "Complexity and scientific modeling." *Foundations of Science* 5, no. 3 (2000): 379-390.
- Evans, Jonathan St BT. *Bias in human reasoning: Causes and consequences*. Lawrence Erlbaum Associates, Inc, 1989.
- Frankish, Keith, and William M. Ramsey, eds. *The Cambridge Handbook of Artificial Intelligence*. Cambridge University Press, 2014.
- Geary, David C., and Kelly J. Huffman. "Brain and cognitive evolution: forms of modularity and functions of mind." *Psychological bulletin* 128, no. 5 (2002): 667.
- Helander, Martin G., ed. *Handbook of human-computer interaction*. Elsevier, 2014.

---

<sup>23</sup> Frankish, Keith, and William M. Ramsey, eds. *The Cambridge Handbook of Artificial Intelligence*. Cambridge University Press, 2014.

- Hoey, Jesse, Tobias Schröder, and Areej Alhothali. "Affect control processes: Intelligent affective interaction using a partially observable Markov decision process." *arXiv preprint arXiv:1306.5279* (2013).
- Jackson, Frank. "Epiphenomenal qualia." *The Philosophical Quarterly* 32, no. 127 (1982): 127-136.
- Kirby, Simon, Tom Griffiths, and Kenny Smith. "Iterated learning and the evolution of language." *Current opinion in neurobiology* 28 (2014): 108-114.
- Millán, José del R., Rüdiger Rupp, Gernot Mueller-Putz, Roderick Murray-Smith, Claudio Giugliemma, Michael Tangermann, Carmen Vidaurre et al. "Combining brain-computer interfaces and assistive technologies: state-of-the-art and challenges." *Frontiers in neuroscience* 4 (2010): 161.
- Negnevitsky, Michael. *Artificial intelligence: a guide to intelligent systems*. Pearson Education, 2005.
- North, Rockstar, Wydawca Rockstar Games, and Wydawca Cenega Poland. "Grand theft auto." *Rockstar Games* (1998).
- Rossen, Brent, Kyle Johnsen, Adeline Deladisma, Scott Lind, and Benjamin Lok. "Virtual humans elicit skin-tone bias consistent with real-world skin-tone biases." In *International Workshop on Intelligent Virtual Agents*, pp. 237-244. Springer Berlin Heidelberg, 2008.
- Schweizer, Paul. "Could There be a Turing Test for Qualia?." *Revisiting Turing and His Test: Comprehensiveness, Qualia, and The Real World*: 41, <http://www.frontiersinai.com/turingfiles/July/07.pdf#page=41> (accessed December 2015)
- Sloman, Aaron, and Ron Chrisley. "Virtual machines and consciousness." *Journal of Consciousness Studies* 10, no. 4-5 (2003): 133-172.
- Terzopoulos, Demetri. "Artificial life for computer graphics." *Communications of the ACM* 42, no. 8 (1999): 32-42.
- Tsagarakis, Nikolaos G., Giorgio Metta, Giulio Sandini, David Vernon, Ricardo Beira, Francesco Becchi, Ludovic Righetti et al. "iCub: the design and realization of an open humanoid platform for cognitive and neuroscience research." *Advanced Robotics* 21, no. 10 (2007): 1151-1175.